1. Performs all three searches

The ontology search module performs all three types of searches in the order of decreasing specificity: gap relationship search, relationship analogues search, and keyword ontology relations search. The results of all three queries are combined and shown to the user in the same order.

1. Ontology Results layout

Duplicate results are removed, with more specific results appearing closer to the top of the results. Each result includes the model name and how it is related to one or more of the keywords.



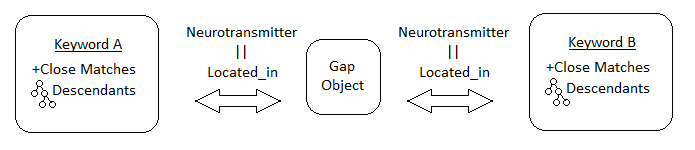
1. Keyword Parsing

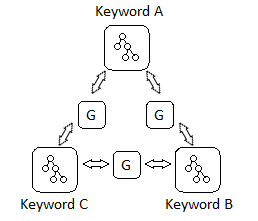
Each keyword or quote phrase is parsed separately, and associates its get\_close\_matches from the ontology entity list, and any of its Is\_part\_of hierarchical descendants. The get\_close\_matches function uses Ratcliff-Metzener (1988) algorithm, while the descendants are found using a recursive (e.g. ?child Is\_part\_of+ ?parent) SPARQL query.

Ratcliff, John W., and David E. Metzener. "Pattern-matching-the gestalt approach." *Dr Dobbs Journal* 13.7 (1988): 46.

1. Gap searches are performed between keyword hierarchies

To find entities satisfying the gap relationship, each keyword or its close matches, and any hierarchical descendants are combined into a set of entity IDs. Then, for each possible unique pair of keywords, an entry of a UNION SPARQL query is generated that combines the results of SPARQL sub-queries that find the gap relationships between the pairs. The pair sub-queries are bi-directional and return any entities that satisfy the gap relationship between the keyword entity ID sets.





1. Fuseki server

To create the ontology, Neurolex ontology relations Id, Located\_in, Neurotransmitter, and Is\_part\_of were loaded into an instance of the Apache Jena Fuseki (<http://jena.apache.org/>) server. The Fuseki server exposes an interface that allows issuing SPARQL queries against its stored ontology via HTTP GET requests and receiving a response in JSON format.

1. Performance tests

To assess the performance of the search functionality, we selected 26 search queries that represented a sample of characteristic terms searched by neuroscientists searching for NeuroML models. We included cases where no results were expected, as well as complex queries multiple keywords. The queries were issued sequentially 20 times each in random order, and the number of milliseconds after issuing the request to receiving a response from the server was measured. The average load times for each query and their standard errors are shown below.

